MAXIMUM LIKELIHOOD ESTIMATION OF THE COVARIANCES OF THE VECTOR MOVING AVERAGE MODELS IN THE TIME AND FREQUENCY DOMAINS

BY

F. AHRABI

TECHNICAL REPORT NO. 36
AUGUST 1978

PREPARED UNDER CONTRACT N00014-75-C-0442
(NR-042-034)
OFFICE OF NAVAL RESEARCH

THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

MAXIMUM LIKELIHOOD ESTIMATION OF THE COVARIANCES OF THE VECTOR MOVING AVERAGE MODELS IN THE TIME AND FREQUENCY DOMAINS

BY

F. AHRABI

Technical Report No. 36

August 1978

PREPARED UNDER CONTRACT N00014-75-C-0442

(NR-042-034)

OFFICE OF NAVAL RESEARCH

Theodore W. Anderson, Project Director

DEPARTMENT OF STATISTICS

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

Maximum Likelihood Estimation of the Covariances of the Vector Moving Average Models in the Time and Frequency Domains

by

F. Ahrabi
Stanford University

Abstract

The vector moving average process is a stationary stochastic process $\{\underset{\sim}{y}_t\}$ satisfying $\underset{\sim}{y}_t = \sum_{i=0}^{q} \underset{\sim}{A}_i \underset{\sim}{\varepsilon}_{t-i}$, where the unobservable process $\{\underset{\sim}{\varepsilon}_t\}$ consists of independently identically distributed random variables. The matrix parameters $\underset{\sim}{\Sigma}^{(s)} = \mathcal{E}\underset{\sim}{y}_t \underset{\sim}{y}'_{t+s}$, $s = 0, 1, \ldots, q$ are estimated from the observations $\underset{\sim}{y}_1, \ldots, \underset{\sim}{y}_T$. The likelihood function is derived under normality and to solve the maximum likelihood equations the Newton-Raphson and Scoring methods are used. The estimation problem is considered in the time and frequency domains. Asymptotic efficiency of the estimates is established.

Key words: Maximum likelihood estimation, vector moving average models, Newton-Raphson and Scoring iterative procedures, Time and Frequency Domains.

Maximum Likelihood Estimation of the Covariances of the Vector Moving Average Models in the Time and Frequency Domains

by

F. Ahrabi
Stanford University

1. Introduction

There have been a number of papers dealing with the estimation of the vector autoregressive moving average (VARMA) models. Hannan (1969, 1970) has considered the problem in the pure moving average case in the frequency domain. Nicholls (1976) has extended this to the case of VARMA models which include exogenous variables. Reinsel (1976) has considered the problem in the time domain. There have been other papers in this area, among them Akaike (1973), Tunnicliffe Wilson (1973), Kashyap (1970), Whittle (1963), and Osborn (1977). In all these papers the parameters of interest are the matrix coefficients of the vectors of observable and unobservable random variables and the common variance covariance matrix of the unobservable random variables.

This paper is concerned with the estimation of vector moving average models, but following Anderson (1975), Parzen (1971), and Clevenson (1970) we take as our parameters the autocovariance matrices of the observable random variables.

There is one paper by Newton (1975) which is primarily concerned with the usual parameterization of the VARMA models. However he derives estimates for the autocovariance matrices in the pure moving average case. His method is to regress the elements of the sample spectral

Acknowledgement. I would like to thank my adviser, Professor T. W. Anderson, under whose supervision and guidance this work was completed.

density, evaluated at a number of equidistant points, on certain trigonometric functions using the method of weighted least squares. The estimate seems to be different from the estimates derived in this paper.

The methods used in this paper are the Newton-Raphson and Scoring Methods, applied to the maximum likelihood equations in the time and frequency domains. The likelihood function is derived under the assumption of normality of the data.

To summarize, Section 2 describes the model and the parameters to be estimated. Sections 3 and 4 deal with the estimation problem in the time and frequency domains respectively. Finally, in Section 5 we derive the limiting average information matrix and show that the estimates proposed are consistent and have the desired limiting multivariate normal distribution, i.e. they are asymptotically efficient.

2. The Model

We have the observations $\underset{\sim}{y}_1, \underset{\sim}{y}_2, \ldots, \underset{\sim}{y}_T$, where

$$(2.1) \qquad \underset{\sim}{y}_t = \underset{p\times 1}{\underset{\sim}{\varepsilon}_t} + \underset{p\times p}{\underset{\sim}{A}_1} \underset{p\times 1}{\underset{\sim}{\varepsilon}_{t-1}} + \ldots + \underset{\sim}{A}_q \underset{\sim}{\varepsilon}_{t-q} .$$

Assumption 1: The $\underset{\sim}{\varepsilon}_t$'s are i.i.d. with mean zero and unknown covariance matrix $\underset{\sim}{V}$.

Assumption 2: The roots of the determinantal equation

$$(2.2) \qquad |\underset{\sim}{I} + \underset{\sim}{A}_1 z + \underset{\sim}{A}_2 z^2 + \ldots + \underset{\sim}{A}_q z^q| = 0$$

lie outside the unit circle.

Note: Assumption 2 enables us to recover the coefficients $\underset{\sim}{A}_1, \underset{\sim}{A}_2, \ldots, \underset{\sim}{A}_q$ from the autocovariance matrices uniquely. The latter are the parameters of interest which are

$$\underset{\sim}{\Sigma}^{(0)} = \mathcal{E}(\underset{\sim}{y}_t \underset{\sim}{y}_t') ,$$

$$\underset{\sim}{\Sigma}^{(s)} = \mathcal{E}(\underset{\sim}{y}_t \underset{\sim}{y}_{t+s}'), \; s=1, 2, \ldots, q .$$

For ease of differentiation of the log likelihood function we should vectorize these matrices where we use the notation

$$\text{Vec } \underset{\sim}{A} = \begin{pmatrix} \underset{\sim}{a}_1 \\ \vdots \\ \underset{\sim}{a}_n \end{pmatrix} ,$$

where

$$\underset{\sim}{A} = (\underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_n)$$

and the $\underset{\sim}{a}_i$'s are column vectors. But we notice that $\underset{\sim}{\Sigma}^{(0)}$ is symmetric and hence should be treated separately. So we vectorize the diagonal and subdiagonal elements of $\underset{\sim}{\Sigma}^{(0)}$ separately. So the parameters are

$$\underset{\sim}{\theta}_0^{(1)} = dg(\underset{\sim}{\Sigma}^{(0)}) = \begin{pmatrix} \sigma_{11}^{(0)} \\ \vdots \\ \sigma_{pp}^{(0)} \end{pmatrix},$$

$$\underset{\sim}{\theta}_0^{(2)} = \widetilde{\text{Vec}}\ \underset{\sim}{\Sigma}^{(0)} ,$$

where the operator $\widetilde{\text{Vec}}$ vectorizes any matrix that it is applied to, ignoring the diagonal and upper diagonal elements of that matrix, e.g.,

$$\widetilde{\text{Vec}} \begin{pmatrix} 1 & 2 & 9 \\ 4 & 3 & 5 \\ 0 & 6 & 7 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 6 \end{pmatrix},$$

and finally

$$\underset{\sim}{\theta}_s = \text{Vec}\ \underset{\sim}{\Sigma}^{(s)},\ s=1,\ 2,\ \ldots,\ q\ .$$

Now we can put all these vectors in a single $\left[qp^2 + \frac{p(p+1)}{2}\right] \times 1$ vector $\underset{\sim}{\theta}$, i.e.

$$\underset{\sim}{\theta}' = (\underset{\sim}{\theta}_0',\ \underset{\sim}{\theta}_1',\ \ldots,\ \underset{\sim}{\theta}_q')$$

where

$$\underset{\sim}{\theta}_0' = (\underset{\sim}{\theta}_0^{'(1)},\ \underset{\sim}{\theta}_0^{'(2)}) .$$

Remarks.

(i) We can find a matrix $\underset{p\times p^2}{\underset{\sim}{B}}$ such that

(2.3) $$dg(\underset{p\times p}{A}) = \underset{\sim}{B}\ \text{Vec}\ \underset{\sim}{A} .$$

$\underset{\sim}{B}$ is obtained from the $p^2 \times p^2$ identity matrix by deleting all the rows except 1st, p+2nd, 2p+3rd, ..., p^2th, i.e.

(2.4)
$$\underset{\sim}{B} = \begin{pmatrix} \underset{\sim}{e}_1' \\ \underset{\sim}{e}_{p+2}' \\ \underset{\sim}{e}_{2p+3}' \\ \vdots \\ \underset{\sim}{e}_{p^2}' \end{pmatrix} ,$$

where

$$\underset{\sim}{I}_{p^2} = (\underset{\sim}{e}_1, \underset{\sim}{e}_2, \ldots, \underset{\sim}{e}_{p^2}) .$$

(ii) In a similar manner we can find a $\frac{p(p-1)}{2} \times p^2$ matrix $\underset{\sim}{C}$ such that

(2.5)
$$\widetilde{\mathrm{Vec}} \underset{p \times p}{\underset{\sim}{A}} = \underset{\sim}{C} \,\mathrm{Vec}\, \underset{\sim}{A} .$$

$\underset{\sim}{C}$ is obtained from $\underset{\sim}{I}_{p^2}$ by deleting the following rows

$$\begin{array}{rrrr} 1, \; p+1, & 2p+1, & \ldots, & (p-1)p+1 \\ p+2, & 2p+2, & \ldots, & (p-1)p+2 \\ & 2p+3, & \ldots, & (p-1)p+3 \\ & & & \vdots \\ & & & (p-1)p+p \; . \end{array}$$

We shall find it convenient to introduce another vector $\underset{\sim}{\tilde{\theta}}$ where

(2.6)
$$\begin{aligned} \underset{\sim}{\tilde{\theta}}' &= (\mathrm{Vec}' \, \underset{\sim}{\Sigma}^{(0)}, \mathrm{Vec}' \, \underset{\sim}{\Sigma}^{(1)}, \ldots, \mathrm{Vec}' \, \underset{\sim}{\Sigma}^{(q)}) \\ &= (\underset{\sim}{\tilde{\theta}}_0', \underset{\sim}{\tilde{\theta}}_1', \ldots, \underset{\sim}{\tilde{\theta}}_q') . \end{aligned}$$

3. Estimation in the Time Domain

3.1. Introduction

We are going to use maximum likelihood estimation and proceed as if the $\boldsymbol{\varepsilon}_t$'s in Section 2 are normally distributed and we shall show later that the resulting estimates have the same limiting covariance matrix irrespective of $\boldsymbol{\varepsilon}_t$'s being normal.

Let

$$\mathbf{y}' = (\mathbf{y}_1', \ldots, \mathbf{y}_T') \ .$$

Then

$$\mathbf{y} \sim N(\mathbf{0}, \boldsymbol{\Sigma}) \ ,$$

where

$$\boldsymbol{\Sigma} = \begin{pmatrix}
\boldsymbol{\Sigma}^{(0)} & \boldsymbol{\Sigma}^{(1)} & \cdots & \boldsymbol{\Sigma}^{(q)} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\
\boldsymbol{\Sigma}'^{(1)} & \boldsymbol{\Sigma}^{(0)} & \boldsymbol{\Sigma}^{(1)} & \cdots & \boldsymbol{\Sigma}^{(q)} & \mathbf{0} & \cdots & \mathbf{0} \\
\vdots & \ddots & \ddots & \ddots & & \ddots & & \\
\boldsymbol{\Sigma}'^{(q)} & & \ddots & \ddots & \ddots & & \ddots & \\
\mathbf{0} & \ddots & & \ddots & \ddots & \ddots & & \boldsymbol{\Sigma}^{(q)} \\
\vdots & & \ddots & & \ddots & \ddots & \ddots & \vdots \\
\mathbf{0} & & & \boldsymbol{\Sigma}'^{(q)} & \cdots & \boldsymbol{\Sigma}'^{(1)} & \boldsymbol{\Sigma}^{(0)} & \boldsymbol{\Sigma}^{(1)}
\end{pmatrix} .$$

Now using Kronecker products and the $T \times T$ matrix

$$\underset{\sim}{L} = \begin{pmatrix} \underset{\sim}{0} & \underset{\sim}{I}_{T-1} \\ 0 & \underset{\sim}{0} \end{pmatrix}$$

similar to the one introduced by Anderson (1975) we can write $\underset{\sim}{\Sigma}$ as

$$(3.1.1) \quad \underset{\sim}{\Sigma} = \underset{\sim}{I}_T \otimes \underset{\sim}{\Sigma}^{(0)} + (\underset{\sim}{L} \otimes \underset{\sim}{\Sigma}^{(1)} + \underset{\sim}{L}' \otimes \underset{\sim}{\Sigma}'^{(1)}) + \ldots + (\underset{\sim}{L}^q \otimes \underset{\sim}{\Sigma}^{(q)} + \underset{\sim}{L}'^q \otimes \underset{\sim}{\Sigma}'^{(q)}) \ .$$

The log likelihood of $\underset{\sim}{y}$ is

$$\log \ell(\underset{\sim}{y}) = -\frac{T}{2} \log 2n - \tfrac{1}{2} \log|\underset{\sim}{\Sigma}| - \tfrac{1}{2} \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} \ .$$

The maximum likelihood estimates are a set of the roots of the equation

$$\frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}} = 0 \ .$$

So we proceed to find the first derivative of the log likelihood, and in doing so we use the fact that

$$\mathcal{E}_{\underset{\sim}{\theta}}\left(\frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}}\right) = 0 \ ,$$

which means

$$(3.1.2) \qquad \frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}} = -\tfrac{1}{2} \frac{\partial}{\partial \underset{\sim}{\theta}} \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} + \tfrac{1}{2} \mathcal{E}\left(\frac{\partial}{\partial \underset{\sim}{\theta}} \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}\right) \ .$$

3.2. The First Derivative of $\log \ell(\underset{\sim}{y})$

As noted above we only need to find $\frac{\partial}{\partial \underset{\sim}{\theta}} \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}$. It is more convenient to find $\frac{\partial}{\partial \underset{\sim}{\theta}} \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}$ and then, noting that

$$(3.2.1)\qquad \frac{\partial}{\partial\sigma_{ij}^{(0)}}\, \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} = -\underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\, \frac{\partial \underset{\sim}{\Sigma}}{\partial\sigma_{ij}^{(0)}}\, \underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}$$

$$= -\underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\,(\underset{\sim}{I}_T \otimes \underset{\sim}{E}_{ij})\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}$$

$$= -\underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\,(\underset{\sim}{I}_T \otimes \underset{\sim}{E}_{ji})\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}$$

$$= \frac{\partial}{\partial\sigma_{ji}^{(0)}}\, \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}\ ,$$

where $\underset{\sim}{E}_{ij}$ is a matrix with 1 for the ij^{th} element and 0's elsewhere, we get

$$(3.2.2)\qquad \frac{\partial}{\partial\tilde{\sigma}_{ij}^{(0)}}\, \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} = 2\, \frac{\partial}{\partial\sigma_{ij}^{(0)}}\, \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}\ ,$$

where $\dfrac{\partial}{\partial\tilde{\sigma}_{ij}^{(0)}}$ indicates that we take the symmetry of $\underset{\sim}{\Sigma}^{(0)}$ into account. The end result is that

$$(3.2.3)\qquad \frac{\partial}{\partial\underset{\sim}{\theta}}\, \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} = \underset{\sim}{G}\, \frac{\partial}{\partial\underset{\sim}{\tilde{\theta}}}\, \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}\ ,$$

where $\underset{\sim}{G}$ is a $\left[qp^2 + \frac{p(p+1)}{2}\right] \times (q+1)p^2$ matrix which can be written as

$$(3.2.4)\qquad \underset{\sim}{G} = \begin{pmatrix} \underset{\sim}{G}_1 \\ \underset{\sim}{G}_2 \\ \underset{\sim}{G}_3 \end{pmatrix} \begin{array}{l} p \\ \frac{p(p-1)}{2} \\ qp^2 \end{array}\ ,$$

and

$$\underset{\sim}{G}_1 = (\underset{\sim}{B},\ \underset{\sim}{0}),\quad \underset{\sim}{G}_2 = (2\underset{\sim}{C},\ \underset{\sim}{0}),\quad \underset{\sim}{G}_3 = (\underset{\sim}{0},\ \underset{\sim}{I})\ ,$$

with $\underset{\sim}{B}$ and $\underset{\sim}{C}$ as in Section 2.

So we proceed to find $\frac{\partial}{\partial \tilde{\theta}} \underset{\sim}{y}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{y}$, using the fact

$$\frac{\partial \underset{\sim}{\Sigma}^{-1}}{\partial x} = -\underset{\sim}{\Sigma}^{-1} \frac{\partial \underset{\sim}{\Sigma}}{\partial x} \underset{\sim}{\Sigma}^{-1} .$$

We shall also find it convenient to express $\underset{\sim}{y}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{y}$ differently using the identity

$$\mathrm{Vec}(\underset{\sim}{A}\underset{\sim}{B}\underset{\sim}{C}) = (\underset{\sim}{C}' \otimes \underset{\sim}{A}) \, \mathrm{Vec} \, \underset{\sim}{B} ,$$

[see Minc and Marcus (1964)] which enables us to write

$$\underset{\sim}{y}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{y} = \mathrm{Vec}(\underset{\sim}{y}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{y}) = (\underset{\sim}{y}' \otimes \underset{\sim}{y}') \, \mathrm{Vec} \, \underset{\sim}{\Sigma}^{-1} .$$

So we only have to differentiate $\mathrm{Vec} \, \underset{\sim}{\Sigma}^{-1}$, but

$$\frac{\partial \, \mathrm{Vec} \, \underset{\sim}{\Sigma}^{-1}}{\partial x} = -\mathrm{Vec}\left(\underset{\sim}{\Sigma}^{-1} \frac{\partial \underset{\sim}{\Sigma}}{\partial x} \underset{\sim}{\Sigma}^{-1}\right) = -(\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1}) \, \mathrm{Vec} \, \frac{\theta \underset{\sim}{\Sigma}}{\partial x} .$$

Now

$$\frac{\partial \underset{\sim}{\Sigma}}{\partial \sigma_{ij}^{(0)}} = \underset{\sim}{I}_T \otimes \underset{\sim}{E}_{ij} , \quad i, j=1, \ldots, p ,$$

$$\frac{\partial \underset{\sim}{\Sigma}}{\partial \sigma_{ij}^{(s)}} = \underset{\sim}{L}^{s} \otimes \underset{\sim}{E}_{ij} + \underset{\sim}{L}'^{s} \otimes \underset{\sim}{E}_{ji}, \quad i, j=1, \ldots, p .$$

So

$$\frac{\partial \, \mathrm{Vec} \, \underset{\sim}{\Sigma}^{-1}}{\partial \sigma_{ij}^{(0)}} = -(\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1}) \, \mathrm{Vec}(\underset{\sim}{I}_T \otimes \underset{\sim}{E}_{ij}), \quad i, j=1, \ldots, p$$

$$= -(\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1}) \, \underset{\sim}{\rho}_{ij}^{(0)} \quad \text{say,}$$

which yields

(3.2.5) $$\frac{\partial \text{ Vec } \underset{\sim}{\Sigma}^{-1}}{\partial \underset{\sim 0}{\tilde{\theta}'}} = -(\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1}) \underset{\sim 0}{E} ,$$

where

(3.2.6) $$\underset{\sim 0}{E} = (\underset{\sim 11}{\rho}^{(0)}, \underset{\sim 21}{\rho}^{(0)}, \ldots, \underset{\sim pp}{\rho}^{(0)}) .$$

Similarly, differentiating w.r.t. $\sigma_{ij}^{(s)}$ we get

$$\frac{\partial \text{ Vec } \underset{\sim}{\Sigma}^{-1}}{\partial \sigma_{ij}^{(s)}} = -(\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1}) \text{ Vec}(\underset{\sim}{L}^{s} \otimes \underset{\sim ij}{E} + \underset{\sim}{L}'^{s} \otimes \underset{\sim ji}{E})$$

$$= -(\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1}) \rho_{ij}^{(s)} , \quad i, j=1, \ldots, p ,$$

which yields

(3.2.7) $$\frac{\partial \text{ Vec } \underset{\sim}{\Sigma}^{-1}}{\partial \underset{\sim s}{\tilde{\theta}'}} = -(\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1}) \underset{\sim s}{E}, \quad s = 1, \ldots, q ,$$

where

$$\underset{\sim s}{E} = (\rho_{11}^{(s)}, \rho_{21}^{(s)}, \ldots, \rho_{pp}^{(s)}) .$$

From (3.2.5) and (3.2.7) we get

(3.2.8) $$\frac{\partial \text{ Vec } \underset{\sim}{\Sigma}^{-1}}{\partial \underset{\sim}{\tilde{\theta}'}} = -(\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1}) \underset{\sim}{E} ,$$

where

(3.2.9) $$\underset{\sim}{E} = (\underset{\sim 0}{E}, \underset{\sim 1}{E}, \ldots, \underset{\sim q}{E}) ,$$

which in turn yields

(3.2.10) $$\frac{\partial \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}}{\partial \underset{\sim}{\tilde{\theta}}'} = -(\underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1})\, \underset{\sim}{E} .$$

Now to complete the computation of $\frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\tilde{\theta}}}$ we have to take the expectation of (3.2.10). We note that for any two vectors $\underset{\sim}{u}$, $\underset{\sim}{v}$ (of the same dimension) we have

$$\mathrm{Vec}(\underset{\sim}{u}\underset{\sim}{v}') = \underset{\sim}{v} \otimes \underset{\sim}{u} .$$

This means

$$(\underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}) = \mathrm{Vec}'(\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}\underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}) .$$

So

$$\begin{aligned} \mathcal{E}(\underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}) &= \mathrm{Vec}'\, \mathcal{E}(\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}\underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}) \\ &= \mathrm{Vec}'\, \underset{\sim}{\Sigma}^{-1} . \end{aligned}$$

Therefore

$$\frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\tilde{\theta}}'} = \tfrac{1}{2}(\underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1} - \mathrm{Vec}'\, \underset{\sim}{\Sigma}^{-1})\, \underset{\sim}{E} ,$$

or

$$\frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\tilde{\theta}}} = \tfrac{1}{2}\, \underset{\sim}{E}'(\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} \otimes \underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} - \mathrm{Vec}\, \underset{\sim}{\Sigma}^{-1}) .$$

Finally

(3.2.11) $$\frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}} = \tfrac{1}{2}\, \underset{\sim}{G}\underset{\sim}{E}'(\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} \otimes \underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} - \mathrm{Vec}\, \underset{\sim}{\Sigma}^{-1}) .$$

<u>Note</u>: We have from above

$$\frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\tilde{\theta}}_i} = \tfrac{1}{2}\, \underset{\sim}{E}_i'(\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} \otimes \underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} - \mathrm{Vec}\, \underset{\sim}{\Sigma}^{-1}) .$$

Now

$$\underset{\sim}{E}_i'(\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} \otimes \underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}) = \begin{pmatrix} \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{A}_{i1}\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} \\ \vdots \\ \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{A}_{ip^2}\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} \end{pmatrix},$$

where Vec $\underset{\sim}{A}_{ir}$ is the r^{th} column of $\underset{\sim}{E}_i$ and $\underset{\sim}{A}_{ir}$ is $Tp \times Tp$. So we get

$$\underset{\sim}{E}_i'(\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} \otimes \underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}) = (\underset{\sim}{I}_{p^2} \otimes \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1})\ \underset{\sim}{A}_i\underset{\sim}{\Sigma}^{-1}y\ ,$$

where

$$\underset{\sim}{A}_i' = (\underset{\sim}{A}_{i1}', \ldots, \underset{\sim}{A}_{ip^2}') \ .$$

This means

$$(3.2.12) \qquad \frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}_i} = \tfrac{1}{2}(\underset{\sim}{I}_{p^2} \otimes \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1})\ \underset{\sim}{A}_i\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} - \tfrac{1}{2}\underset{\sim}{E}_i' \text{ Vec } \underset{\sim}{\Sigma}^{-1} \ .$$

We shall use this latter form for $\dfrac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}_i}$ when finding the second derivative of $\log \ell(\underset{\sim}{y})$.

3.3. The Numerical Approximations

The equation

$$\frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}} = 0$$

is clearly nonlinear and cannot be solved explicitly. So we will use numerical approximations to get asymptotically efficient estimates. The methods we are going to use are the Newton-Raphson method and the

Scoring method. In both methods we need an initial consistent estimate (of order $T^{\frac{1}{2}}$) of $\underset{\sim}{\theta}$, call it $\hat{\underset{\sim}{\theta}}_{(0)}$, then the Newton-Raphson method consists of solving the following system of linear equations for $\hat{\underset{\sim}{\theta}}_{(1)}$,

$$(3.3.1) \qquad - \left.\frac{\partial^2 \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}\, \partial \underset{\sim}{\theta}'}\right|_{\underset{\sim}{\theta}=\hat{\underset{\sim}{\theta}}_{(0)}} (\hat{\underset{\sim}{\theta}}_{(1)} - \hat{\underset{\sim}{\theta}}_{(0)}) = \left.\frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}}\right|_{\underset{\sim}{\theta}=\hat{\underset{\sim}{\theta}}_{(0)}} .$$

The Scoring method is the same as above with $\mathcal{E}\, \frac{\partial^2 \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}\, \partial \underset{\sim}{\theta}'}$ replacing $\frac{\partial^2 \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}\, \partial \underset{\sim}{\theta}'}$, i.e. we have the equations

$$(3.3.2) \qquad -\mathcal{E}\left.\left(\frac{\partial^2 \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}\, \partial \underset{\sim}{\theta}'}\right)\right|_{\underset{\sim}{\theta}=\hat{\underset{\sim}{\theta}}_{(0)}} (\hat{\underset{\sim}{\theta}}_{(1)} - \hat{\underset{\sim}{\theta}}_{(0)}) = \left.\frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}}\right|_{\underset{\sim}{\theta}=\hat{\underset{\sim}{\theta}}_{(0)}} .$$

To find $\hat{\underset{\sim}{\theta}}_{(0)}$ we estimate the covariance matrices by their sample analogues, so

$$(3.3.3) \qquad \hat{\underset{\sim}{\Sigma}}^{(s)}_{(0)} = \frac{1}{T-s} \sum_{t=1}^{T-s} \underset{\sim}{y}_t \underset{\sim}{y}'_{t+s} , \quad s = 0, 1, \ldots, q .$$

3.4. The Scoring Method

To arrive at the linear equations for this method we need to find $\mathcal{E}\left(\frac{\partial^2 \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}\, \partial \underset{\sim}{\theta}'}\right)$, but we know that

$$\mathcal{E}\left(\frac{\partial^2 \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}\, \partial \underset{\sim}{\theta}'}\right) = -\mathcal{E}\left(\frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}} \cdot \frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}'}\right)$$

and from (3.2.11)

$$(3.4.1) \qquad \mathcal{E}\left(\frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}} \cdot \frac{\partial \log \ell(\underset{\sim}{y})}{\partial \underset{\sim}{\theta}'}\right) =$$

$$\tfrac{1}{4}\underset{\sim\sim}{GE}'\, \mathcal{E}[(\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} \otimes \underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} - \operatorname{Vec} \underset{\sim}{\Sigma}^{-1})(\underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1} - \operatorname{Vec}' \underset{\sim}{\Sigma}^{-1})]\underset{\sim\sim}{EG}' .$$

Now

$$\underset{\sim}{z} \equiv \underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} \sim N(0,\underset{\sim}{\Sigma}^{-1})$$

and we need

$$\mathcal{E}(\underset{\sim}{z} \otimes \underset{\sim}{z} - \text{Vec } \underset{\sim}{\Sigma}^{-1})(\underset{\sim}{z}' \otimes \underset{\sim}{z}' - \text{Vec}' \underset{\sim}{\Sigma}^{-1})$$

$$= \mathcal{E}(\underset{\sim}{z}\underset{\sim}{z}' \otimes \underset{\sim}{z}\underset{\sim}{z}') - \text{Vec } \underset{\sim}{\Sigma}^{-1} \text{ Vec}' \underset{\sim}{\Sigma}^{-1}$$

since $\mathcal{E}(\underset{\sim}{z} \otimes \underset{\sim}{z}) = \text{Vec } \underset{\sim}{\Sigma}^{-1}$.

So, suppose $\underset{\sim}{u} \sim N(\underset{\sim}{0}, \underset{n\times n}{\underset{\sim}{D}})$, we want to find $\mathcal{E}(\underset{\sim}{u}\underset{\sim}{u}' \otimes \underset{\sim}{u}\underset{\sim}{u}')$. The ij^{th} block of $\underset{\sim}{u}\underset{\sim}{u}' \otimes \underset{\sim}{u}\underset{\sim}{u}'$ is

$$u_i u_j \underset{\sim}{u}\underset{\sim}{u}' .$$

To find $\mathcal{E}(u_i u_j \underset{\sim}{u}\underset{\sim}{u}')$ we use the result (Anderson (1958, p. 39)).

$$\mathcal{E}(u_i u_j u_r u_s) = d_{ij}d_{rs} + d_{ir}d_{js} + d_{is}d_{jr} ,$$

which yields

$$\mathcal{E}(u_i u_j \underset{\sim}{u}\underset{\sim}{u}') = d_{ij} \underset{\sim}{D} + \underset{\sim}{d}_i\underset{\sim}{d}_j' + \underset{\sim}{d}_j\underset{\sim}{d}_i' ,$$

where

$$\underset{\sim}{D} = (d_1, \ldots, d_n) .$$

From this we can get

$$\mathcal{E}(\underset{\sim}{u}\underset{\sim}{u}' \otimes \underset{\sim}{u}\underset{\sim}{u}') = \underset{\sim}{D} \otimes \underset{\sim}{D} + \begin{pmatrix} \underset{\sim}{d}_1 \\ \vdots \\ \underset{\sim}{d}_n \end{pmatrix} (\underset{\sim}{d}_1', \ldots, \underset{\sim}{d}_n')$$

$$+ \begin{pmatrix} \underset{\sim}{d}_1\underset{\sim}{d}_1' & \underset{\sim}{d}_2\underset{\sim}{d}_1' & \cdots & \underset{\sim}{d}_n\underset{\sim}{d}_1' \\ \vdots & & & \\ \underset{\sim}{d}_1\underset{\sim}{d}_n' & \underset{\sim}{d}_2\underset{\sim}{d}_n' & \cdots & \underset{\sim}{d}_n\underset{\sim}{d}_n' \end{pmatrix} .$$

Now, the last matrix was shown by Magnus and Neudecker (1977) to be equal to

$$\underset{\sim}{K}_n(\underset{\sim}{D} \otimes \underset{\sim}{D}) ,$$

where

$$(3.4.2) \qquad \underset{\sim}{K}_n = \begin{pmatrix} \underset{\sim}{E}_{11}^{(n)} & \underset{\sim}{E}_{21}^{(n)} & \cdots & \underset{\sim}{E}_{n1}^{(n)} \\ \vdots & & & \\ \underset{\sim}{E}_{1n}^{(n)} & \underset{\sim}{E}_{2n}^{(n)} & \cdots & \underset{\sim}{E}_{nn}^{(n)} \end{pmatrix}$$

and $\underset{\sim}{E}_{ij}^{(n)}$'s are $n\times n$ matrices as defined before. So

$$\mathcal{E}(\underset{\sim}{u}\underset{\sim}{u}' \otimes \underset{\sim}{u}\underset{\sim}{u}') = (\underset{\sim}{I}_{n^2} + \underset{\sim}{K}_n)(\underset{\sim}{D} \otimes \underset{\sim}{D}) + \text{Vec } \underset{\sim}{D} \text{ Vec}' \underset{\sim}{D} .$$

Now for our problem

$$\underset{\sim}{u} = \underset{\sim}{z},\ \underset{\sim}{D} = \underset{\sim}{\Sigma}^{-1},\ n = Tp$$

which means

$$(3.4.3) \quad \mathcal{E}(\underset{\sim}{z}\underset{\sim}{z}' \otimes \underset{\sim}{z}\underset{\sim}{z}') = (\underset{\sim}{I}_{T^2p^2} + \underset{\sim}{K}_{Tp})(\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1}) + \text{Vec } \underset{\sim}{\Sigma}^{-1} \text{ Vec}' \underset{\sim}{\Sigma}^{-1} .$$

Finally we get the average information matrix as $1/T$ times

$$\mathcal{E}\left(\frac{\partial \log \ell}{\partial \underset{\sim}{\theta}} \cdot \frac{\partial \log \ell}{\partial \underset{\sim}{\theta}'}\right) = \tfrac{1}{4}\, \underset{\sim}{G}\underset{\sim}{E}' (\underset{\sim}{I}_{T^2p^2} + \underset{\sim}{K}_{Tp})(\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1})\, \underset{\sim}{E}\underset{\sim}{G}' .$$

So the linear equations for the scoring method are

$$\text{(3.4.5)} \qquad \underset{\sim\sim}{GE}'(\underset{\sim}{I}_{T^2p^2} + \underset{\sim}{K}_{Tp})(\hat{\underset{\sim}{\Sigma}}^{-1}_{(0)} \otimes \hat{\underset{\sim}{\Sigma}}^{-1}_{(0)})\ \underset{\sim\sim}{EG}'(\hat{\underset{\sim}{\theta}}_{(1)} - \hat{\underset{\sim}{\theta}}_{(0)})$$

$$= 2\ \underset{\sim\sim}{GE}'(\hat{\underset{\sim}{\Sigma}}^{-1}_{(0)}\underset{\sim}{y} \otimes \hat{\underset{\sim}{\Sigma}}^{-1}_{(0)}\underset{\sim}{y} - \text{Vec}\ \hat{\underset{\sim}{\Sigma}}^{-1}_{(0)}) \ .$$

Once we get $\hat{\underset{\sim}{\theta}}_{(1)}$, we could use that as $\hat{\underset{\sim}{\theta}}_{(0)}$ in (3.4.5) and get a second iterate $\hat{\underset{\sim}{\theta}}_{(2)}$, but for large samples this is not necessary.

3.5. The Newton-Raphson Method

3.5.1. Preliminaries

To write down the linear equations for this method we need the second derivative of the log likelihood function. To derive the latter we first derive $\dfrac{\partial^2}{\partial\tilde{\underset{\sim}{\theta}}\ \partial\tilde{\underset{\sim}{\theta}}'} \log \ell(\underset{\sim}{y})$, using the form (3.2.12) for $\dfrac{\partial \log \ell}{\partial\tilde{\underset{\sim}{\theta}}}$ and then use

$$\text{(3.5.1.1)} \qquad \frac{\partial^2 \log \ell}{\partial\underset{\sim}{\theta}\ \partial\underset{\sim}{\theta}'} = \underset{\sim}{G}\ \frac{\partial^2 \log \ell}{\partial\tilde{\underset{\sim}{\theta}}\ \partial\tilde{\underset{\sim}{\theta}}'}\ \underset{\sim}{G}' \ .$$

3.5.2. The Derivation of $\dfrac{\partial^2 \log \ell}{\partial\tilde{\underset{\sim}{\theta}}\ \partial\tilde{\underset{\sim}{\theta}}'}$.

As in (3.2.12) we have

$$\frac{\partial \log \ell}{\partial\tilde{\underset{\sim}{\theta}}_i} = \tfrac{1}{2}(\underset{\sim}{I}_{p^2} \otimes \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1})\ \underset{\sim}{A}_i\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y} - \tfrac{1}{2}\ \underset{\sim}{E}_i'\ \text{Vec}\ \underset{\sim}{\Sigma}^{-1} \ ,$$

$$i = 0, 1, \ldots, q \ .$$

We have the derivative of the second term w.r.t. $\tilde{\underset{\sim}{\theta}}$ as

$$\text{(3.5.2.1)} \qquad \tfrac{1}{2}\ \underset{\sim}{E}_i(\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1})\ \underset{\sim}{E} \ ,$$

using (3.2.8). It remains to find the derivative of the first term. To do this we shall find the derivative w.r.t. $\tilde{\theta}_j$ for $i \geq j$ and using symmetry will complete the derivation. So we let

$$\tau_i \equiv \tau_i(\tilde{\theta}) = \tfrac{1}{2}(I_{p^2} \otimes y'\Sigma^{-1})\, A_i \Sigma^{-1} y, \quad i = 0, 1, \ldots, q$$

and

$$\tau' = (\tau_0', \ldots, \tau_q') \,.$$

Now

$$\frac{\partial \tau_i}{\partial \sigma_{uv}^{(0)}} = -\tfrac{1}{2}[I_{p^2} \otimes y'\Sigma^{-1}(I_T \otimes E_{uv})\Sigma^{-1}]\, A_i \Sigma^{-1} y$$

$$- \tfrac{1}{2}(I_{p^2} \otimes y'\Sigma^{-1})\, A_i \Sigma^{-1}(I_T \otimes E_{uv})\Sigma^{-1} y \,.$$

Now using $(A \otimes B)(C \otimes D) = (AC \otimes BD)$, we can factor out $(I_{p^2} \otimes y'\Sigma^{-1})$ to the left and also can factor out $\Sigma^{-1} y$ to the right, which results in

$$\frac{\partial \tau_i}{\partial \tilde{\theta}_{uv}^{(0)}} = -\tfrac{1}{2}(I_{p^2} \otimes y'\Sigma^{-1})\, C_{i0,uv}\, \Sigma^{-1} y \,,$$

where

$$C_{i0,uv} = [I_{p^2} \otimes (I_T \otimes E_{uv})\Sigma^{-1}]\, A_i + A_i \Sigma^{-1}(I_T \otimes E_{uv}) \,.$$

This leads to

$$\frac{\partial \tau_i}{\partial \tilde{\theta}_0'} = \left(\frac{\partial \tau_i}{\partial \sigma_{11}^{(0)}}, \frac{\partial \tau_i}{\partial \sigma_{21}^{(0)}}, \ldots, \frac{\partial \tau_i}{\partial \sigma_{pp}^{(0)}} \right)$$

$$= -\tfrac{1}{2}(I_{p^2} \otimes y'\Sigma^{-1})\, (C_{i0,11}\, \Sigma^{-1} y, \ldots, C_{i0,pp}\, \Sigma^{-1} y)$$

$$= -\tfrac{1}{2}(I_{p^2} \otimes y'\Sigma^{-1})\, C_{i0}(I_{p^2} \otimes \Sigma^{-1} y) \,,$$

where

$$\underset{\sim}{C}_{i0} = (\underset{\sim}{C}_{i0,11}, \underset{\sim}{C}_{i0,21}, \ldots, \underset{\sim}{C}_{i0,pp}) , \quad i = 0, 1, \ldots, q .$$

Similarly

$$\frac{\partial \underset{\sim}{\tau}_i}{\partial \underset{\sim}{\tilde{\theta}}_j'} = -\tfrac{1}{2}(\underset{\sim}{I}_{p^2} \otimes \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}) \underset{\sim}{C}_{ij} (\underset{\sim}{I}_{p^2} \otimes \underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}) ,$$

where

$$\underset{\sim}{C}_{ij} = (\underset{\sim}{C}_{ij,11}, \underset{\sim}{C}_{ij,21}, \ldots, \underset{\sim}{C}_{ij,pp})$$

and

$$\underset{\sim}{C}_{ij,uv} = [\underset{\sim}{I}_{p^2} \otimes (\underset{\sim}{L}^j \otimes \underset{\sim}{E}_{uv} + \underset{\sim}{L}'^j \underset{\sim}{E}_{vu})\underset{\sim}{\Sigma}^{-1}] \underset{\sim}{A}_i + \underset{\sim}{A}_i\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{L}^j \otimes \underset{\sim}{E}_{uv} + \underset{\sim}{L}'^j \underset{\sim}{E}_{vu}) ,$$

$$j \leq i, \; i = 0, 1, \ldots, q ,$$

$$u,v = 1, 2, \ldots p .$$

Finally for $j \geq i$ using symmetry we have

$$\frac{\partial \underset{\sim}{\tau}_i}{\partial \underset{\sim}{\tilde{\theta}}_j'} = -\tfrac{1}{2}(\underset{\sim}{I}_{p^2} \otimes \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}) \underset{\sim}{C}_{ij}' (\underset{\sim}{I}_{p^2} \otimes \underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}) .$$

Now, for $j \geq i$ define

$$\underset{\sim}{C}_{ij} \equiv \underset{\sim}{C}_{ji}'$$

and define

$$\underset{\sim}{C} = (\underset{\sim}{C}_{ij}) , \quad i,j = 0, 1, \ldots, q ,$$

then from above we have

$$(3.5.2.2)\qquad \frac{\partial \tilde{\tau}}{\partial \tilde{\theta}'} = -\tfrac{1}{2}\left((\underset{\sim}{I}_{p^2} \otimes \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1})\, \underset{\sim}{C}_{ij}(\underset{\sim}{I}_{p^2} \otimes \underset{\sim}{\Sigma}^{-1}\underset{\sim}{y})\right)_{i,j\,=\,0,\,1,\,\ldots,\,q}$$

$$= -\tfrac{1}{2}[\underset{\sim}{I}_{q+1} \otimes (\underset{\sim}{I}_{p^2} \otimes \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1})]\, \underset{\sim}{C}[\underset{\sim}{I}_{q+1} \otimes (\underset{\sim}{I}_{p^2} \otimes \underset{\sim}{\Sigma}^{-1}\underset{\sim}{y})]\ .$$

Now notice that

$$\underset{\sim}{I}_n \otimes (\underset{\sim}{I}_m \otimes \underset{\sim}{A}) \equiv \underset{\sim}{I}_{mn} \otimes \underset{\sim}{A}\ .$$

So (3.5.2.2) becomes

$$\frac{\partial \tilde{\tau}}{\partial \tilde{\theta}'} = -\tfrac{1}{2}(\underset{\sim}{I}_{p^2(q+1)} \otimes \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1})\, \underset{\sim}{C}(\underset{\sim}{I}_{p^2(q+1)} \otimes \underset{\sim}{\Sigma}^{-1}\underset{\sim}{y})\ .$$

Finally from (3.5.2.1) we have

$$(3.5.2.3)\qquad \frac{\partial^2 \log \ell}{\partial \tilde{\theta}\ \partial \tilde{\theta}'} = -\tfrac{1}{2}(\underset{\sim}{I}_{p^2(q+1)} \otimes \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1})\, \underset{\sim}{C}(\underset{\sim}{I}_{p^2(q+1)} \otimes \underset{\sim}{\Sigma}^{-1}\underset{\sim}{y})$$

$$+ \tfrac{1}{2}\, \underset{\sim}{E}'(\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1})\, \underset{\sim}{E}\ .$$

3.5.3. The Equations

From (3.5.2.3) using (3.5.1.1) we have

$$\frac{\partial^2 \log \ell}{\partial \tilde{\theta}\ \partial \tilde{\theta}'} = -\tfrac{1}{2}\, \underset{\sim}{G}(\underset{\sim}{I}_{p^2(q+1)} \otimes \underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1})\, \underset{\sim}{C}(\underset{\sim}{I}_{p^2(q+1)} \otimes \underset{\sim}{\Sigma}^{-1}\underset{\sim}{y})\, \underset{\sim}{G}'$$

$$+ \tfrac{1}{2}\, \underset{\sim}{G}\underset{\sim}{E}'(\underset{\sim}{\Sigma}^{-1} \otimes \underset{\sim}{\Sigma}^{-1})\, \underset{\sim}{E}\underset{\sim}{G}'\ .$$

So the equations for the Newton-Raphson method are

$$[\underset{\sim}{G}(\underset{\sim}{I}_{p^2(q+1)} \otimes \underset{\sim}{y}'\hat{\underset{\sim}{\Sigma}}^{-1}_{(0)})\, \hat{\underset{\sim}{C}}_{(0)}(\underset{\sim}{I}_{p^2(q+1)} \otimes \hat{\underset{\sim}{\Sigma}}^{-1}_{(0)}\underset{\sim}{y}) - \underset{\sim}{G}\underset{\sim}{E}'(\hat{\underset{\sim}{\Sigma}}^{-1}_{(0)} \otimes \hat{\underset{\sim}{\Sigma}}^{-1}_{(0)})\, \underset{\sim}{E}\underset{\sim}{G}']$$

$$\times\ (\hat{\underset{\sim}{\theta}}_{(1)} - \hat{\underset{\sim}{\theta}}_{(0)}) = \underset{\sim}{G}\underset{\sim}{E}'(\hat{\underset{\sim}{\Sigma}}^{-1}_{(0)}\underset{\sim}{y} \otimes \hat{\underset{\sim}{\Sigma}}^{-1}_{(0)}\underset{\sim}{y} - \mathrm{Vec}\ \hat{\underset{\sim}{\Sigma}}^{-1}_{(0)})\ .$$

3.6. Summary

In the preceding sections we have derived the linear equations for the estimates, using the Newton-Rapnson and Scoring methods. The two resulting estimates are asymptotically equivalent, but at this point it is not clear which one is easier to compute. To write down the equations we first have to invert the $Tp \times Tp$ block Toeplitz matrix $\hat{\underset{\sim}{\Sigma}}_{(0)}$ and once we have the equations, we need to find the best method (computationally easiest) to solve them. These problems will be considered in future work.

To compare the computational problems with that of Reinsel (1976), we see that in the latter a matrix $\hat{\underset{\sim}{G}}$ which is essentially of the same form and size as $\hat{\underset{\sim}{\Sigma}}_{(0)}$ has to be inverted $(\underset{\sim}{G} = \sum_{i=0}^{q} \underset{\sim}{A}_i \otimes \underset{\sim}{L}^i)$. But once $\hat{\underset{\sim}{G}}^{-1}$ is computed, Reinsel has shown that the solutions to his equations are the same as some generalized least square estimates.

4. Estimation in the Frequency Domain

4.1. Introduction

For a stationary process $\{\underset{\sim}{z}_t,\ t = 0, \pm 1, \ldots\}$ with covariances $\underset{\sim}{D}_s = \mathcal{E}(\underset{\sim}{z}_t \underset{\sim}{z}'_{t+s})$, $s = 0, \pm 1, \ldots$, the spectral density matrix $\underset{\sim}{f}$ is defined as

(4.1.1) $$\underset{\sim}{f}(\lambda) = \frac{1}{2\pi} \sum_{s=-\infty}^{\infty} \underset{\sim}{D}_s\, e^{-is\lambda} .$$

The covariances can be recovered from $\underset{\sim}{f}(\cdot)$ via

$$\underset{\sim}{D}_s = \int_{-\pi}^{\pi} \underset{\sim}{f}(\lambda)\, e^{is\lambda}\, d\lambda .$$

The sample analogue of the spectral density, the periodogram, is defined as

(4.1.2) $$\underset{\sim}{I}(\lambda) = \frac{1}{2\pi} \sum_{-(T-1)}^{T-1} \hat{\underset{\sim}{D}}_s\, e^{-is\lambda} ,$$

where $\hat{\underset{\sim}{D}}_s$ is the sample analogue of $\underset{\sim}{D}_s$, more precisely

$$\hat{\underset{\sim}{D}}_s = \frac{1}{T} \sum_{t=1}^{T-s} \underset{\sim}{y}_t \underset{\sim}{y}'_{t+s} .$$

We can also represent $\underset{\sim}{I}(\lambda)$ in terms of the discrete Fourier transforms

(4.1.3) $$\underset{\sim}{I}(\lambda) = \underset{\sim}{w}(\lambda)\, \underset{\sim}{w}^*(\lambda) ,$$

where

(4.1.4) $$\underset{\sim}{w}(\lambda) = \frac{1}{\sqrt{2\pi T}} \sum_{n=1}^{T} \underset{\sim}{z}_n\, e^{in\lambda} .$$

For a fuller treatment see Anderson (1971).

If the process $\{\underset{\sim}{z}_t,\ t = 0, \pm 1, \ldots\}$ is Gaussian then the log likelihood function may be approximated by

$$(4.1.5) \qquad -\tfrac{1}{2} \log |\underset{\sim}{\Gamma}| - \tfrac{1}{2} \sum_t \operatorname{tr}[\underset{\sim}{f}^{-1}(\lambda_t)\, \underset{\sim}{I}(\lambda_t)] \; ,$$

where

$$\underset{\sim}{\Gamma} = \operatorname{Cov}(\underset{\sim}{z}_1, \ldots, \underset{\sim}{z}_T) \; ,$$

$$\lambda_t = \frac{2\pi t}{T}, \; t = 0, 1, \ldots, T-1 \; .$$

Whittle (1953, 1961) suggested this for the case $p=1$ and Dunsmuir and Hannan (1976) showed that this leads to efficient estimates for general p.

For our problem $\underset{\sim}{y} \equiv \underset{\sim}{z}$, $\underset{\sim}{D}_s \equiv \underset{\sim}{\Sigma}^{(s)}$, $\underset{\sim}{\Gamma} = \underset{\sim}{\Sigma}$, and we only have a finite number of nonzero covariances so

$$(4.1.6) \qquad \underset{\sim}{f}(\lambda) = \frac{1}{2\pi} \sum_{-q}^{q} \underset{\sim}{\Sigma}^{(s)} e^{-i\lambda s}$$

and

$$\underset{\sim}{I}(\lambda) = \frac{1}{2\pi} \sum_{-(T-1)}^{T-1} \hat{\underset{\sim}{D}}_s \, e^{-i\lambda s} \; .$$

So the log likelihood can be approximated by

$$(4.1.7) \qquad \log \ell \approx -\tfrac{1}{2} \log |\underset{\sim}{\Sigma}| - \tfrac{1}{2} \sum_t \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t) \; ,$$

where $\underset{\sim}{f}_t \equiv \underset{\sim}{f}(\lambda_t)$, $\underset{\sim}{I}_t = \underset{\sim}{I}(\lambda_t)$. We shall use the same approximation methods as in the time domain and will also use

$$(4.1.8) \qquad \frac{\partial \log \ell}{\partial \underset{\sim}{\theta}} = \underset{\sim}{G} \frac{\partial \log \ell}{\partial \underset{\sim}{\tilde{\theta}}} \; ,$$

$$(4.1.9) \qquad \mathcal{E}\left(\frac{\partial \log \ell}{\partial \underset{\sim}{\tilde{\theta}}}\right) = 0 \; .$$

4.2. The Derivation of $\frac{\partial \log \ell}{\partial \tilde{\underset{\sim}{\theta}}}$

Using (4.1.7) and (4.1.9) we get

$$\frac{\partial \log \ell}{\partial \underset{\sim}{\tilde{\theta}}} = -\tfrac{1}{2} \sum_t \left[\frac{\partial}{\partial \underset{\sim}{\tilde{\theta}}} \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t) - \mathcal{E} \frac{\partial}{\partial \underset{\sim}{\tilde{\theta}}} \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t) \right] . \tag{4.2.1}$$

Now

$$\frac{\partial}{\partial \underset{\sim}{\tilde{\theta}}} \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t) = -\operatorname{tr}(\underset{\sim}{f}_t^{-1} \frac{\partial \underset{\sim}{f}_t}{\partial x} \underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t) .$$

Differentiating (4.1.6) yields

$$\frac{\partial \underset{\sim}{f}_t}{\partial \sigma_{uv}^{(0)}} = \frac{1}{2\pi} \underset{\sim}{E}_{uv} ,$$

$$\frac{\partial \underset{\sim}{f}_t}{\partial \sigma_{uv}^{(s)}} = \frac{1}{2\pi} [e^{-i\lambda_t s} \underset{\sim}{E}_{uv} + e^{i\lambda_t s} \underset{\sim}{E}_{vu}] , \quad u,v = 1, \ldots, p .$$

So

$$\begin{aligned} \frac{\partial}{\partial \sigma_{uv}^{(0)}} \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t) &= -\frac{1}{2\pi} \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{E}_{uv} \underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t) \\ &= -\frac{1}{2\pi} \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{e}_u \underset{\sim}{e}_v' \underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t) \\ &= -\frac{1}{2\pi} \underset{\sim}{e}_v' \underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t \underset{\sim}{f}_t^{-1} \underset{\sim}{e}_u \\ &= -\frac{1}{2\pi} (\underset{\sim}{h}_t)_{vu}, \quad u,v = 1, \ldots, p , \end{aligned}$$

where

$$\underset{\sim}{h}_t = \underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t \underset{\sim}{f}_t^{-1} , \tag{4.2.2}$$

and we have used the fact that $\underset{\sim}{E}_{uv} = \underset{\sim}{e}_u \underset{\sim}{e}_v'$, where $\underset{\sim}{e}_u$, $\underset{\sim}{e}_v$ are the u^{th} and v^{th} column of the $p \times p$ identity matrix. Now we can easily see

$$(4.2.3) \qquad \frac{\partial}{\partial \underset{\sim}{\tilde{\theta}}_0} \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t) = -\frac{1}{2\pi} \operatorname{Vec}(\underset{\sim}{h}_t') .$$

Similarly

$$\frac{\partial}{\partial \sigma_{uv}^{(s)}} \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t) = -\frac{1}{2\pi} \operatorname{tr}\left[\underset{\sim}{f}_t^{-1}(e^{-i\lambda_t s} \underset{\sim}{E}_{uv} + e^{i\lambda_t s} \underset{\sim}{E}_{vu}) \underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t\right]$$

$$= -\frac{1}{2\pi}\left[e^{-i\lambda_t s} \underset{\sim}{h}_t' + e^{i\lambda_t s} \underset{\sim}{h}_t\right]_{uv}, \quad u,v = 1, \ldots, p ,$$

which yields

$$(4.2.4) \qquad \frac{\partial}{\partial \underset{\sim}{\tilde{\theta}}_s} \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t) = -\frac{1}{2\pi} \operatorname{Vec}(e^{i\lambda_t s} \underset{\sim}{h}_t + e^{-i\lambda_t s} \underset{\sim}{h}_t'), \quad s = 1, \ldots, q .$$

To complete the derivation of $\frac{\partial \log \underset{\sim}{\ell}}{\partial \underset{\sim}{\theta}}$ we have to take the expectation of (4.2.3) and (4.2.4). This yields

$$\mathcal{E}\left[\frac{\partial}{\partial \underset{\sim}{\tilde{\theta}}_0} \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)\right] = -\frac{1}{2\pi} \operatorname{Vec}(\underset{\sim}{f}_t'^{-1}) ,$$

$$\mathcal{E}\left[\frac{\partial}{\partial \underset{\sim}{\tilde{\theta}}_s} \operatorname{tr}(\underset{\sim}{f}_t^{-1} \underset{\sim}{I}_t)\right] = -\frac{1}{2\pi} \operatorname{Vec}\left[e^{i\lambda_t s} \underset{\sim}{f}_t^{-1} + e^{-i\lambda_t s} \underset{\sim}{f}_t'^{-1}\right] ,$$

since $\mathcal{E}\underset{\sim}{h}_t \sim \underset{\sim}{f}_t^{-1}$, which follows from $\mathcal{E}(\underset{\sim}{I}_t) = \underset{\sim}{f}_t + O(T^{-1})$. Now let

$$\underset{\sim}{\ell}_t = \underset{\sim}{h}_t - \underset{\sim}{f}_t^{-1} ,$$

then using (4.2.1) we have

(4.2.5) $$\frac{\partial \log \ell}{\partial \tilde{\theta}_0} = \frac{1}{4\pi} \sum_t \text{Vec } \ell'_t \ ,$$

(4.2.6) $$\frac{\partial \log \ell}{\partial \tilde{\theta}_s} = \frac{1}{4\pi} \sum_t \text{Vec}\left[e^{i\lambda_t s} \ell_t + e^{-i\lambda_t s} \ell'_t \right] \ , \quad s = 1, \ldots, q \ .$$

Now, from (4.1.6) it is obvious that $f' = \bar{f}$ or $f = f^*$, where $*$ indicates "conjugate transpose", i.e. f is Hermitian. Also $I' = \bar{I}$, which leads to $h' = \bar{h}$ and $\ell' = \bar{\ell}$. We can use this to simplify (4.2.6) as follows

$$\frac{\partial \log \ell}{\partial \tilde{\theta}_s} = \frac{1}{4\pi} \sum_t \text{Vec}(e^{i\lambda_t s} \ell_t) + \overline{\frac{1}{4\pi} \sum_t \text{Vec}(e^{i\lambda_t s} \ell_t)}$$

$$= \frac{1}{2\pi} \sum_t \text{Vec}(e^{i\lambda_t s} \ell_t) \ ,$$

because the first sum is real. The reason for this is that

$$\overline{e^{i\lambda_t}} = e^{-i\lambda_t} = e^{-i \frac{2\pi t}{T}} = e^{i2\pi \frac{(T-t)}{T}} = e^{i\lambda_{T-t}} \ .$$

This means that for any real function $\eta(\cdot)$

$$\sum_{t=0}^{T-1} \eta(e^{i\lambda_t}) = \eta(0) + \sum_{t=0}^{\frac{T-1}{2}} \eta(e^{i\lambda_t}) + \eta(\overline{e^{i\lambda_t}}) \ ,$$

for T odd, and

$$\sum_{t=0}^{T-1} \eta(e^{i\lambda_t}) = \eta(0) + \sum_{t=0}^{\frac{T-1}{2}} \eta(e^{i\lambda_t}) + \eta(\overline{e^{i\lambda_t}}) + \eta(-1) \ ,$$

for T even. So $\sum_{t=0}^{T-1} \eta(e^{i\lambda_t})$ is real. The same argument allows us to rewrite (4.2.5) as

$$\frac{\partial \log \ell}{\partial \tilde{\underset{\sim}{\theta}}} = \frac{1}{4\pi} \sum_t \text{Vec } \underset{\sim}{\ell}_t \ .$$

Finally we can conclude

$$\frac{\partial \log \ell}{\partial \tilde{\underset{\sim}{\theta}}_0} = \frac{1}{2\pi} \sum_t \underset{\sim}{J}_t \text{ Vec } \underset{\sim}{\ell}_t \ , \tag{4.2.7}$$

where

$$\underset{\sim}{J}_t' = (\tfrac{1}{2} \underset{\sim}{I}_{p^2}, e^{i\lambda t} \underset{\sim}{I}_{p^2}, \ldots, e^{qi\lambda_t} \underset{\sim}{I}_{p^2}) \ .$$

There is an alternative form for (4.2.7) which will be more useful in deriving the second derivative of $\log \ell$. It is obtained by noticing that we can find a matrix $\underset{\sim}{M}$, such that

$$\text{Vec } \underset{\sim}{\ell}_t = \underset{\sim}{M} \text{ Vec } \underset{\sim}{\ell}_t' \ ,$$

which then enables us to rewrite (4.2.6) as

$$\frac{\partial \log \ell}{\partial \tilde{\underset{\sim}{\theta}}_s} = \frac{1}{4\pi} \sum_t (e^{i\lambda_t s} \underset{\sim}{I}_{p^2} + e^{-i\lambda_t s} \underset{\sim}{M}) \text{ Vec } \underset{\sim}{\ell}_t \ . \tag{4.2.8}$$

It is easily verified that in fact

$$\underset{\sim}{M} = \underset{\sim}{K}_p \ ,$$

where $\underset{\sim}{K}_p$ was defined in (3.4.2). So the alternative form for the first derivative of $\log \ell$ is

$$\frac{\partial \log \ell}{\partial \tilde{\underset{\sim}{\theta}}} = \frac{1}{4\pi} \sum_t \underset{\sim}{H}_t \text{ Vec } \underset{\sim}{\ell}_t \ , \tag{4.2.9}$$

where

$$(4.2.10)\quad \underset{\sim}{H}_t' \equiv \underset{\sim}{H}'(\lambda_t) = (\underset{\sim}{I}_{p^2},\ e^{i\lambda_t}\underset{\sim}{I}_{p^2} + e^{-i\lambda_t}\underset{\sim}{K}_p,\ \ldots,\ e^{qi\lambda_t}\underset{\sim}{I}_{p^2} + e^{-pi\lambda_t}\underset{\sim}{K}_p) \ .$$

4.3. The Second Derivative of log ℓ

Using (4.2.9) we have

$$(4.3.1)\qquad \frac{\partial^2 \log \ell}{\partial\underset{\sim}{\tilde\theta}\ \partial\underset{\sim}{\tilde\theta}'} = \frac{1}{4\pi}\sum_t \underset{\sim}{H}_t \frac{\partial \operatorname{Vec} \underset{\sim}{\ell}_t}{\partial\underset{\sim}{\tilde\theta}'} \ .$$

Now

$$\underset{\sim}{\ell}_t = \underset{\sim}{f}_t^{-1}\underset{\sim}{I}_t\underset{\sim}{f}_t^{-1} - \underset{\sim}{f}_t^{-1} ,$$

which yields

$$\frac{\partial\underset{\sim}{\ell}_t}{\partial x} = -\underset{\sim}{f}_t^{-1}\frac{\partial\underset{\sim}{f}_t}{\partial x}\underset{\sim}{f}_t^{-1}\underset{\sim}{I}_t\underset{\sim}{f}_t^{-1} - \underset{\sim}{f}_t^{-1}\underset{\sim}{I}_t\underset{\sim}{f}_t^{-1}\frac{\partial\underset{\sim}{f}_t}{\partial x}\underset{\sim}{f}_t^{-1} + \underset{\sim}{f}_t^{-1}\frac{\partial\underset{\sim}{f}_t}{\partial x}\underset{\sim}{f}_t^{-1} ,$$

which in turn yields

$$(4.3.2)\quad \frac{\partial \operatorname{Vec}\underset{\sim}{\ell}_t}{\partial x} = [(\underset{\sim}{f}_t^{'-1}\otimes\underset{\sim}{f}_t^{-1}) - (\underset{\sim}{f}_t^{'-1}\otimes\underset{\sim}{h}_t) - (\underset{\sim}{h}_t'\otimes\underset{\sim}{f}_t^{-1})]\operatorname{Vec}\frac{\partial\underset{\sim}{f}_t}{\partial x}$$

$$= \underset{\sim}{\phi}_t \operatorname{Vec}\frac{\partial\underset{\sim}{f}_t}{\partial x} ,$$

say. Now for $x = \sigma_{uv}^{(0)}$ we get

$$\frac{\partial \operatorname{Vec}\underset{\sim}{\ell}_t}{\partial\sigma_{uv}^{(0)}} = \frac{1}{2\pi}\underset{\sim}{\phi}_t\underset{\sim}{e}_{uv} ,$$

where

$$\underset{\sim}{e}_{uv} = \operatorname{Vec}\underset{\sim}{E}_{uv},\quad u,v = 1, \ldots, p \ .$$

So

$$(4.3.3) \qquad \frac{\partial \operatorname{Vec} \underset{\sim}{\ell}_t}{\partial \underset{\sim}{\tilde{\theta}}_0'} = \frac{1}{2\pi} \underset{\sim}{\phi}_t \, (\underset{\sim}{e}_{11}, \underset{\sim}{e}_{21}, \ldots, \underset{\sim}{e}_{pp})$$

$$= \frac{1}{2\pi} \underset{\sim}{\phi}_t \, \underset{\sim}{I}_{p^2} = \frac{1}{2\pi} \underset{\sim}{\phi}_t \, .$$

Similarly

$$\frac{\partial \operatorname{Vec} \underset{\sim}{\ell}_t}{\partial \sigma_{uv}^{(s)}} = \frac{1}{2\pi} \underset{\sim}{\phi}_t \, (e^{-i\lambda_t s} \underset{\sim}{e}_{uv} + e^{i\lambda_t s} \underset{\sim}{e}_{vu}) \; , \quad u,v = 1, \ldots, p \, .$$

Now

$$\underset{\sim}{e}_{vu} = \operatorname{Vec} \underset{\sim}{E}_{vu} = \operatorname{Vec} \underset{\sim}{E}_{uv}' = \underset{\sim}{K}_p \operatorname{Vec} \underset{\sim}{E}_{uv} = \underset{\sim}{K}_p \, \underset{\sim}{e}_{uv} \; ,$$

which means

$$\frac{\partial \operatorname{Vec} \underset{\sim}{\ell}_t}{\partial \sigma_{uv}^{(s)}} = \frac{1}{2\pi} \underset{\sim}{\phi}_t \, (e^{-i\lambda t_s} \underset{\sim}{I}_{p^2} + e^{i\lambda_t s} \underset{\sim}{K}_p) \, \underset{\sim}{e}_{uv} \; , \quad u,v = 1, \ldots, p \; , \quad s = 1, \ldots, q \, .$$

Finally

$$(4.3.4) \qquad \frac{\partial \operatorname{Vec} \underset{\sim}{\ell}_t}{\partial \underset{\sim}{\tilde{\theta}}_s'} = \frac{1}{2\pi} \underset{\sim}{\phi}_t \, (e^{-i\lambda_t s} \underset{\sim}{I}_{p^2} + e^{i\lambda_t s} \underset{\sim}{K}_p) \; , \quad s = 1, \ldots, q \, .$$

So

$$(4.3.5) \qquad \frac{\partial \operatorname{Vec} \underset{\sim}{\ell}_t}{\partial \underset{\sim}{\tilde{\theta}}'} = \frac{1}{2\pi} \underset{\sim}{\phi}_t (\underset{\sim}{I}_{p^2}, \; e^{-i\lambda_t} \underset{\sim}{I}_{p^2} + e^{i\lambda_t} \underset{\sim}{K}_p, \; \ldots, \; e^{-qi\lambda_t} \underset{\sim}{I}_{p^2} + e^{qi\lambda_t} \underset{\sim}{K}_p \,)$$

$$= \frac{1}{2\pi} \underset{\sim}{\phi}_t \, \underset{\sim}{H}_t^* \, .$$

Now we get the second derivative from (4.3.1)

$$\frac{\partial^2 \log \ell}{\partial \tilde{\underset{\sim}{\theta}}\, \partial \tilde{\underset{\sim}{\theta}}'} = \frac{1}{8\pi^2} \sum_t \underset{\sim}{H}_t\, \underset{\sim}{\phi}_t\, \underset{\sim}{H}_t^* \ . \tag{4.3.6}$$

4.4. The Newton-Raphson Method

As in the time domain we will use $\hat{\underset{\sim}{\Sigma}}^{(s)}_{(0)}$'s as initial estimates of the covariances, and so

$$\hat{\underset{\sim}{f}}_{t(0)} = \frac{1}{2\pi} \sum_{s=-q}^{q} e^{-i\lambda_t s}\, \hat{\underset{\sim}{\Sigma}}^{(s)}_{(0)}$$

will be the initial estimate of $\underset{\sim}{f}_t$. Accordingly we form $\hat{\underset{\sim}{\ell}}_{t(0)}$ and $\hat{\underset{\sim}{\phi}}_{t(0)}$,

$$\hat{\underset{\sim}{\ell}}_{t(0)} = \hat{\underset{\sim}{f}}^{-1}_{t(0)}\, \underset{\sim}{I}_t\, \hat{\underset{\sim}{f}}^{-1}_{t(0)} - \hat{\underset{\sim}{f}}^{-1}_{t(0)} \ ,$$

$$\begin{aligned}\hat{\underset{\sim}{\phi}}_{t(0)} = {} & (\hat{\underset{\sim}{f}}'^{-1}_{t(0)} \otimes \hat{\underset{\sim}{f}}^{-1}_{t(0)}) - (\hat{\underset{\sim}{f}}'^{-1}_{t(0)} \otimes \hat{\underset{\sim}{f}}^{-1}_{t(0)}\, \underset{\sim}{I}_t\, \hat{\underset{\sim}{f}}^{-1}_{t(0)}) \\ & - (\hat{\underset{\sim}{f}}'^{-1}_{t(0)}\, \underset{\sim}{I}'_t\, \hat{\underset{\sim}{f}}'^{-1}_{t(0)} \otimes \hat{\underset{\sim}{f}}^{-1}_{t(0)}) \ .\end{aligned}$$

Using (4.2.7) we have

$$\frac{\partial \log \ell}{\partial \underset{\sim}{\theta}} = \underset{\sim}{G}\, \frac{\partial \log \ell}{\partial \tilde{\underset{\sim}{\theta}}} = \frac{1}{4\pi} \sum_t \underset{\sim}{G}\underset{\sim}{J}_t \operatorname{Vec} \underset{\sim}{\ell}_t \ . \tag{4.4.1}$$

Similarly from (4.3.6) we get

$$\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\theta}\, \partial \underset{\sim}{\theta}'} = \frac{1}{8\pi^2} \sum_t \underset{\sim}{G}\underset{\sim}{H}_t\, \underset{\sim}{\phi}_t\, \underset{\sim}{H}_t^*\, \underset{\sim}{G}' \ . \tag{4.4.2}$$

So the linear equations for the Newton-Raphson method are

$$(4.4.3)\quad (-\sum_t \underset{\sim\sim}{GH}_t\ \underset{\sim}{\hat{\phi}}_{t(0)}\ \underset{\sim}{H}_t^*\underset{\sim}{G}')\ (\underset{\sim}{\hat{\theta}}_{(1)}-\underset{\sim}{\hat{\theta}}_{(0)}) = 2\sum_t \underset{\sim\sim}{GJ}_t\ \text{Vec}\ \underset{\sim}{\hat{\ell}}_{t(0)}\ .$$

4.5. The Scoring Method

As mentioned earlier, to get the linear equations for this method we replace $\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\theta}\, \partial \underset{\sim}{\theta}'}$ by its expectation in the Newton-Raphson method. The latter is

$$\mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\theta}\, \partial \underset{\sim}{\theta}'}\right) = \frac{1}{8\pi^2}\sum_t \underset{\sim\sim}{GH}_t\ \mathcal{E}(\underset{\sim}{\phi}_t)\ \underset{\sim}{H}_t^*\underset{\sim}{G}'\ .$$

Now

$$\mathcal{E}(\underset{\sim}{I}_t) \sim \underset{\sim}{f}_t\ ,$$

which leads to

$$\mathcal{E}(\underset{\sim}{h}_t) = \mathcal{E}(\underset{\sim}{f}_t^{-1}\ \underset{\sim}{I}_t\ \underset{\sim}{f}_t^{-1}) \sim \underset{\sim}{f}_t^{-1}\ ,$$

which in turn leads to

$$\mathcal{E}(\underset{\sim}{\phi}_t) \sim -(\underset{\sim}{f}_t'^{-1} \otimes \underset{\sim}{f}_t^{-1})$$

So

$$\mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial \underset{\sim}{\theta}\, \partial \underset{\sim}{\theta}'}\right) = -\frac{1}{8\pi^2}\sum_t \underset{\sim\sim}{GH}_t(\underset{\sim}{f}_t'^{-1} \times \underset{\sim}{f}_t^{-1})\underset{\sim}{H}_t^*\underset{\sim}{G}'\ ,$$

ignoring the terms of order T^{-1}. Finally the linear equations for this method are

$$[\sum_t \underset{\sim}{G}_t\underset{\sim}{H}_t(\underset{\sim}{f}_{t(0)}'^{-1} \otimes \underset{\sim}{f}_{t(0)}^{-1})\underset{\sim}{H}_t^*\underset{\sim}{G}']\,(\underset{\sim}{\hat{\theta}}_{(1)}-\underset{\sim}{\hat{\theta}}_{(0)}) = 2\sum_t \underset{\sim\sim}{GJ}_t\ \text{Vec}\ \underset{\sim}{\ell}_{t(0)}\ .$$

4.5. Remarks

As in the time domain, there are computational problems to be considered in setting up and solving the equations derived in the preceding sections. In this case the main problem in setting up the equations is the inversion of $p \times p$ matrices $\hat{\underset{\sim}{f}}_t(0)$, $t = 1, \ldots, T-1$. And again we have to find the best way to solve the resulting equations. It seems that the computation of the estimates in the time domain is easier than that in the frequency domain.

Comparing with Nicholls (1976) and Anderson (1978) which deal with the estimation of the coefficients $\underset{\sim}{A}_i$, $i = 1, \ldots, q$, we see that the main problem of inversion of $\hat{\underset{\sim}{f}}_t(0)$'s is also present in these papers.

5. Asymptotic Properties

The four estimates proposed in this paper are asymptotically equivalent and we shall show that they are efficient, i.e.

$$\sqrt{T}\ (\hat{\theta}_{(1)}-\theta) \xrightarrow{\mathcal{L}} N(0,\ \mathcal{J}^{-1}(\theta))\ ,$$

where $\mathcal{J}(\theta)$ is the limiting average informtion matrix and "$\xrightarrow{\mathcal{L}}$" indicates convergence in distribution.

To find $\mathcal{J}(\theta)$, by definition we have

$$(5.1)\qquad \mathcal{J}(\theta) = \lim_{T\to\infty} -\frac{1}{T}\,\mathcal{E}\left(\frac{\partial^2 \log \ell}{\partial\theta\ \partial\theta'}\right)$$

$$= \lim_{T\to\infty}\frac{1}{T}\quad \frac{1}{8\pi^2}\sum_t GH_t(f_t'^{-1}\otimes f_t^{-1})H_t^*G'\ ,$$

$$= \frac{1}{16\pi^3}\int_0^{2\pi} GH(f'^{-1}\otimes f^{-1})H^*G'\ d\lambda\ ,$$

where the argument λ is omitted from H and f.

The four estimates are obtained from equations like

$$(5.2)\qquad \hat{\mathcal{J}}(\hat{\theta}_{(0)})\ (\hat{\theta}_{(1)}-\hat{\theta}_{(0)}) = \frac{1}{T}\left.\frac{\partial \log \ell}{\partial\theta}\right|_{\theta=\hat{\theta}_{(0)}}\ ,$$

where $\hat{\mathcal{J}}(\hat{\theta}_{(0)})$ is a consistent estimate of $\mathcal{J}(\theta)$. We can rewrite (5.2) as

$$(5.3)\qquad \hat{\mathcal{J}}(\hat{\theta}_{(0)})\ \sqrt{T}\ (\hat{\theta}_{(1)}-\theta) = \hat{\mathcal{J}}(\hat{\theta}_{(0)})\ \sqrt{T}\ (\hat{\theta}_{(0)}-\theta) + \frac{1}{\sqrt{T}}\left.\frac{\partial \log \ell}{\partial\theta}\right|_{\theta=\hat{\theta}_{(0)}}\ ,$$

where θ is the true parameter. Now

$$(5.4)\qquad \frac{1}{\sqrt{T}}\frac{\partial \log \ell}{\partial\theta} = \frac{1}{\sqrt{T}}\left.\frac{\partial \log \ell}{\partial\theta}\right|_{\theta=\hat{\theta}_{(0)}} + \frac{1}{\sqrt{T}}\left.\frac{\partial^2 \log \ell}{\partial\theta\ \partial\theta'}\right|_{\theta=\theta^+}(\theta-\hat{\theta}_{(0)})\ ,$$

where $|\underset{\sim}{\theta}-\underset{\sim}{\theta}^{+}| \leq |\underset{\sim}{\theta}-\hat{\underset{\sim}{\theta}}_{(0)}|$. Now (5.3) can be rewritten using (5.4)

$$\text{(5.5)} \qquad \hat{\underset{\sim}{\mathcal{I}}}(\hat{\underset{\sim}{\theta}}_{(0)})\ \sqrt{T}\ (\hat{\underset{\sim}{\theta}}_{(1)}-\underset{\sim}{\theta}) = \left[\hat{\underset{\sim}{\mathcal{I}}}(\hat{\underset{\sim}{\theta}}_{(0)}) + \frac{1}{T}\frac{\partial^2 \log \ell}{\partial\underset{\sim}{\theta}\ \partial\underset{\sim}{\theta}'}\bigg|_{\underset{\sim}{\theta}=\underset{\sim}{\theta}^{+}}\right]$$

$$\sqrt{T}\ (\hat{\underset{\sim}{\theta}}_{(0)}-\underset{\sim}{\theta}) + \frac{1}{\sqrt{T}}\frac{\partial \log \ell}{\partial\underset{\sim}{\theta}} .$$

Now noticing that

$$\frac{1}{T}\frac{\partial^2 \log \ell}{\partial\underset{\sim}{\theta}\ \partial\underset{\sim}{\theta}'}\bigg|_{\underset{\sim}{\theta}=\underset{\sim}{\theta}^{+}} \overset{\mathcal{L}}{\to} -\underset{\sim}{\mathcal{I}}(\underset{\sim}{\theta})$$

and $\sqrt{T}(\hat{\underset{\sim}{\theta}}_{(0)}-\underset{\sim}{\theta})$ is bounded in probability, we see that (5.5) is (asymptotically) equivalent to

$$\text{(5.6)} \qquad \sqrt{T}\ (\hat{\underset{\sim}{\theta}}_{(1)}-\underset{\sim}{\theta}) = \underset{\sim}{\mathcal{I}}^{-1}(\underset{\sim}{\theta})\ \frac{1}{\sqrt{T}}\frac{\partial \log \ell}{\partial\underset{\sim}{\theta}} .$$

<u>Theorem</u>. Under (2.1) and Assumptions 1 and 2 of Section 2,

$$\sqrt{T}\ (\hat{\underset{\sim}{\theta}}_{(1)}-\underset{\sim}{\theta}) \overset{\mathcal{L}}{\to} N(\underset{\sim}{0},\ \underset{\sim}{\mathcal{I}}^{-1}(\underset{\sim}{\theta})) ,$$

where $\hat{\underset{\sim}{\theta}}_{(1)}$ is any one of the four estimates derived in this paper.

<u>Proof</u>. Using (5.6), it suffices to show

$$\frac{1}{\sqrt{T}}\frac{\partial \log \ell}{\partial\underset{\sim}{\theta}} \overset{\mathcal{L}}{\to} N(\underset{\sim}{0},\cdot) .$$

Let

$$\underset{\sim}{\xi} = \begin{pmatrix} \mathrm{Vec}(\underset{\sim}{A}_1,\ \ldots,\ \underset{\sim}{A}_q) \\ \mathrm{dg}\ \underset{\sim}{V} \\ \widetilde{\mathrm{Vec}}\ \underset{\sim}{V} \end{pmatrix} ,$$

where $\underset{\sim}{A}_i$'s and $\underset{\sim}{V}$ were defined in Section 2. Now

$$\frac{\partial \log \ell}{\partial \xi_i} = \sum_j \frac{\partial \log \ell}{\partial \theta_j} \cdot \frac{\partial \theta_J}{\partial \xi_i} ,$$

which means

$$\frac{\partial \log \ell}{\partial \underset{\sim}{\xi}} = \frac{\partial \underset{\sim}{\theta}'}{\partial \underset{\sim}{\xi}} \frac{\partial \log \ell}{\partial \underset{\sim}{\theta}} .$$

It follows from Assumption 2 of Section 2 that $\frac{\partial \underset{\sim}{\theta}'}{\partial \underset{\sim}{\xi}}$ is nonsingular, which means

$$\frac{1}{\sqrt{T}} \frac{\partial \log \ell}{\partial \underset{\sim}{\theta}} = \left(\frac{\partial \underset{\sim}{\theta}'}{\partial \underset{\sim}{\xi}}\right)^{-1} \frac{1}{\sqrt{T}} \frac{\partial \log \ell}{\partial \underset{\sim}{\xi}} . \tag{5.7}$$

But it has been shown by Nicholls (1976) and Reinsel (1976) that

$$\sqrt{T}\ (\hat{\underset{\sim}{\xi}} - \underset{\sim}{\xi}) \overset{\mathcal{L}}{\to} N(\underset{\sim}{0}, \cdot) ,$$

which is the same as

$$\frac{1}{\sqrt{T}} \frac{\partial \log \ell}{\partial \underset{\sim}{\xi}} \overset{\mathcal{L}}{\to} N(\underset{\sim}{0}, \cdot) .$$

So (5.7) gives us

$$\frac{1}{\sqrt{T}} \frac{\partial \log \ell}{\partial \underset{\sim}{\theta}} \overset{\mathcal{L}}{\to} N(\underset{\sim}{0}, \cdot) . \tag{5.8}$$

The limiting covariance matrix in (5.8) is obviously $\underset{\sim}{\mathcal{I}}(\underset{\sim}{\theta})$, so

$$\sqrt{T}\ (\hat{\underset{\sim}{\theta}}_{(1)} - \underset{\sim}{\theta}) \overset{\mathcal{L}}{\to} N(\underset{\sim}{0}, \underset{\sim}{\mathcal{I}}^{-1}(\underset{\sim}{\theta})) . \quad \text{Q.E.D.}$$

References

Akaike, Hirotugu (1973), "Maximum likelihood identification of Gaussian autoregressive moving average model," Biometrika, Vol. 60, pp. 255-265.

Anderson, T. W. (1958), An Introduction to Multivariate Statistical Analysis, John Wiley & Sons, Inc., New York.

Anderson, T. W. (1971), The Statistical Analysis of Time Series, John Wiley & Sons,Inc., New York.

Anderson, T. W. (1975), "Maximum likelihood estimation of parameters of autoregressive processes with moving average residuals and other covariance matrices with linear structure," Annals of Statistics, Vol. 3, pp. 1283-1304.

Anderson, T. W. (1978), "Maximum likelihood estimation for vector autoregressive moving average models," Technical Report No. 35, Department of Statistics, Stanford University, Stanford, CA.

Clevenson, M. Lawrence (1970), "Asymptotically efficient estimates of the parameters of a moving average time series," Technical Report No. 15, Department of Statistics, Stanford University, Stanford, CA.

Dunsmuir, W. and Hannan, E. J. (1976), "Vector linear time series models," Advances in Applied Probability, Vol. 8, pp. 339-364.

Hannan, E. J. (1969), "The estimation of mixed moving average autoregressive systems," Biometrika, Vol. 56, pp. 579-593.

Hannan, E. J. (1970), Multiple Time Series, John Wiley & Sons, Inc., New York.

Kashyap, R. L. (1970), "Maximum likelihood identification of stochastic linear systems," I.E.E.E. Transactions on Automatic Control, Vol. AC-15, pp. 25-34.

Magnus, Jan and Neudecker, H. (1977), "The commutation matrix: Some theorems and applications," Report AE3/77, Institute voor Actuariaat en Econometrie, Universiteit van Amsterdam, Amsterdam.

Minc, Henryk and Marcus, Marvin (1964), A Survey of Matrix Theory and Matrix Inequalities, Prindle, Weber & Schmidt, Boston.

Newton, Howard Joseph (1975), "The efficient estimation of stationary multiple time series mixed models: Theory and algorithms," Technical Report No. 33, Statistical Science Division, State University of New York at Buffalo.

Nicholls, D. F. (1976), "The efficient estimation of vector linear time series models," Biometrika, Vol. 64, pp. 85-90.

Osborn, Denise R. (1977), "Exact and approximate maximum likelihood estimators for vector moving average processes," Journal of the Royal Statistical Society, Series B, Vol. 39, pp. 114-118.

Parzen, Emanuel (1971), "Efficient estimation of stationary time series mixed schemes," Bulletin of the International Statistical Institute, Vol. 44, pp. 315-319.

Reinsel, Gregory C. (1976), "Maximum likelihood estimation of vector autoregressive moving average models," Technical Report No. 117, Department of Statistics, Carnegie-Mellon University, Pittsburgh, PA.

Tunnicliffe Wilson, G. (1973), "The estimation of parameters in multivariate time series models," Journal of the Royal Statistical Society, Series B, Vol. 35, pp. 76-85.

Whittle, P. (1953), "Estimation and information in stationary time series," Arkiv för Matematik, Vol. 2, pp. 423-434.

Whittle, P. (1961), "Gaussian estimation in stationary time series," Bulletin of the International Statistical Institute, Vol. 33, pp. 1-26.

Whittle, P. (1963), "On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix," Biometrika, Vol. 50, pp. 129-134.

TECHNICAL REPORTS

OFFICE OF NAVAL RESEARCH CONTRACT N00014-67-A-0112-0030 (NR-042-034)

1. "Confidence Limits for the Expected Value of an Arbitrary Bounded Random Variable with a Continuous Distribution Function," T. W. Anderson, October 1, 1969.

2. "Efficient Estimation of Regression Coefficients in Time Series," T. W. Anderson, October 1, 1970.

3. "Determining the Appropriate Sample Size for Confidence Limits for a Proportion," T. W. Anderson and H. Burstein, October 15, 1970.

4. "Some General Results on Time-Ordered Classification," D. V. Hinkley, July 30, 1971.

5. "Tests for Randomness of Directions against Equatorial and Bimodal Alternatives," T. W. Anderson and M. A. Stephens, August 30, 1971.

6. "Estimation of Covariance Matrices with Linear Structure and Moving Average Processes of Finite Order," T. W. Anderson, October 29, 1971.

7. "The Stationarity of an Estimated Autoregressive Process," T. W. Anderson, November 15, 1971.

8. "On the Inverse of Some Covariance Matrices of Toeplitz Type," Raul Pedro Mentz, July 12, 1972.

9. "An Asymptotic Expansion of the Distribution of "Studentized" Classification Statistics," T. W. Anderson, September 10, 1972.

10. "Asymptotic Evaluation of the Probabilities of Misclassification by Linear Discriminant Functions," T. W. Anderson, September 28, 1972.

11. "Population Mixing Models and Clustering Algorithms," Stanley L. Sclove, February 1, 1973.

12. "Asymptotic Properties and Computation of Maximum Likelihood Estimates in the Mixed Model of the Analysis of Variance," John James Miller, November 21, 1973.

13. "Maximum Likelihood Estimation in the Birth-and-Death Process," Niels Keiding, November 28, 1973.

14. "Random Orthogonal Set Functions and Stochastic Models for the Gravity Potential of the Earth," Steffen L. Lauritzen, December 27, 1973.

15. "Maximum Likelihood Estimation of Parameters of an Autoregressive Process with Moving Average Residuals and Other Covariance Matrices with Linear Structure," T. W. Anderson, December, 1973.

16. "Note on a Case-Study in Box-Jenkins Seasonal Forecasting of Time Series," Steffen L. Lauritzen, April, 1974.

17. "General Exponential Models for Discrete Observations," Steffen L. Lauritzen, May, 1974.

18. "On the Interrelationships among Sufficiency, Total Sufficiency and Some Related Concepts," Steffen L. Lauritzen, June, 1974.

19. "Statistical Inference for Multiply Truncated Power Series Distributions," T. Cacoullos, September 30, 1974.

Office of Naval Research Contract N00014-75-C-0442 (NR-042-034)

20. "Estimation by Maximum Likelihood in Autoregressive Moving Average Models in the Time and Frequency Domains," T. W. Anderson, June 1975.

21. "Asymptotic Properties of Some Estimators in Moving Average Models," Raul Pedro Mentz, September 8, 1975.

22. "On a Spectral Estimate Obtained by an Autoregressive Model Fitting," Mituaki Huzii, February 1976.

23. "Estimating Means when Some Observations are Classified by Linear Discriminant Function," Chien-Pai Han, April 1976.

24. "Panels and Time Series Analysis: Markov Chains and Autoregressive Processes," T. W. Anderson, July 1976.

25. "Repeated Measurements on Autoregressive Processes," T. W. Anderson, September 1976.

26. "The Recurrence Classification of Risk and Storage Processes," J. Michael Harrison and Sidney I. Resnick, September 1976.

27. "The Generalized Variance of a Stationary Autoregressive Process," T. W. Anderson and Raul P.Mentz, October 1976.

28. "Estimation of the Parameters of Finite Location and Scale Mixtures," Javad Behboodian, October 1976.

29. "Identification of Parameters by the Distribution of a Maximum Random Variable," T. W. Anderson and S.G. Ghurye, November 1976.

30. "Discrimination Between Stationary Guassian Processes, Large Sample Results," Will Gersch, January 1977.

31. "Principal Components in the Nonnormal Case: The Test for Sphericity," Christine M. Waternaux, October 1977.

32. "Nonnegative Definiteness of the Estimated Dispersion Matrix in a Multivariate Linear Model," F. Pukelsheim and George P.H. Styan, May 1978.

33. "Canonical Correlations with Respect to a Complex Structure," Steen A. Andersson, July 1978.

34. "An Extremal Problem for Positive Definite Matrices," T.W. Anderson and I. Olkin, July 1978.

35. " Maximum likelihood Estimation for Vector Autoregressive Moving Average Models," T. W. Anderson, July 1978.

36. "Maximum likelihood Estimation of the Covariances of the Vector Moving Average Models in the Time and Frequency Domains," F. Ahrabi, August 1978.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 36	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) MAXIMUM LIKELIHOOD ESTIMATION OF THE COVARIANCES OF THE VECTOR MOVING AVERAGE MODELS IN THE TIME AND FREQUENCY DOMAINS		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) F. AHRABI		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0442
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, California		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (NR-042-034)
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 436 Arlington, Virginia 22217		12. REPORT DATE AUGUST 1978
		13. NUMBER OF PAGES 36
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Maximum likelihood estimation, vector moving average models, Newton-Raphson and Scoring iterative procedures, Time and Frequency Domains.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

SEE REVERSE SIDE

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT.

The vector moving average process is a stationary stochastic process $\{y_t\}$ satisfying $y_t \; \sum_{i=0}^{q} A_i \varepsilon_{t-i}$, where the unobservable process $\{\varepsilon_t\}$ consists of independently identically distributed random variables. The matrix parameters $\Sigma^{(s)} = \mathcal{E} y_t y_{t+s}'$, $s = 0, 1, \ldots, q$ are estimated from the observations $y_1, \ldots, y_T$. The likelihood function is derived under normality and to solve the maximum likelihood equations the Newton-Raphson and Scoring methods are used. The estimation problem is considered in the time and frequency domains. Asymptotic efficiency of the estimates is established.